Data Spark Consulting project for ProMed

Final Presentation

Haochen Song, Kangyan Zhao, Valentin Courgeau,
Valeria Cortez Vaca Diez and Tobey Agbo-ola

Executive Summary

Modelling conditional probability of (re-) emerging infectious diseases

Findings:

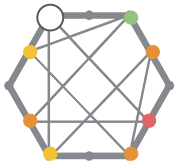
- **Probit regression** most appropriate model
- Some countries are more likely to infect others: **wealthy and well connected countries**
- Some countries more vulnerable to be infected by others: **developing countries and Schengen countries**

Recommendations:

- Use results to **identify most infectious countries and more vulnerable countries**
- **Focus resources** on such countries
- Use country factors **make prediction on possible changes** to improve certain areas of development

ProMED
INTERNATIONAL SOCIETY
FOR INFECTOUS DISEASES

Dataset containing 15 years worth of alerts (+30k) on ~140 diseases from across the world



Model the likelihood that a country infects another for a specific disease

Used country factors, time difference and connection factors as regressors

Agenda

Mid-term review

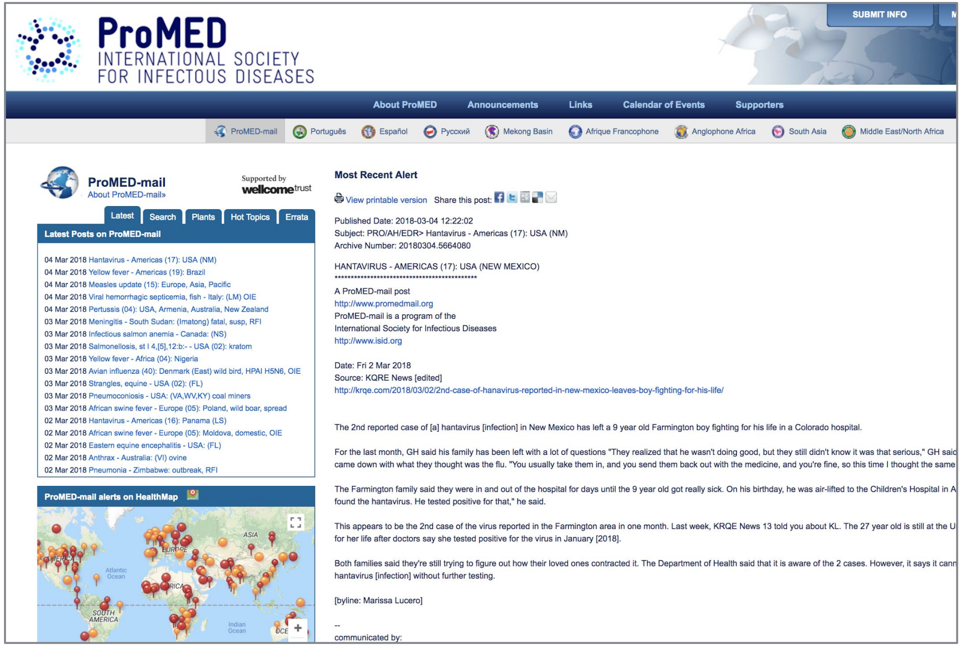
1) Client introduction
2) Dataset
3) Project goal
4) Methodology
5) Results
6) Conclusions and recommendations

Client introduction

Client: ProMED

Internet-based reporting system to monitor emerging diseases

- Non-profit organisation founded in 1999
- Provides early warning of outbreaks of (re-) emerging diseases **to prevent epidemic transmission**
- Sources of information include **media reports, official reports, online summaries, local observers**, and others
- Informs the international infectious disease community, including **scientists, physicians, epidemiologists and public health professionals**



Client: Main Issue

Without a group of data analysts, ProMED's ability to gain insights from free-text reports is very restricted

Cannot analyse the **magnitude of an outbreak**

Cannot visualise and understand **how diseases are connected**

Free-text makes it hard to extract information

No team of specialised data analysts to make use of their data

Dataset

Dataset example

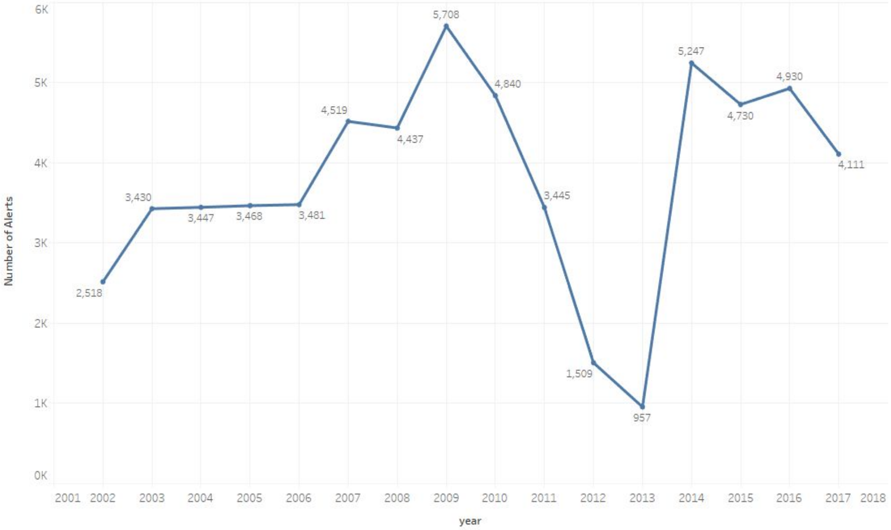
Dataset with few columns. Most important data points were country, date, specie, disease and alert content

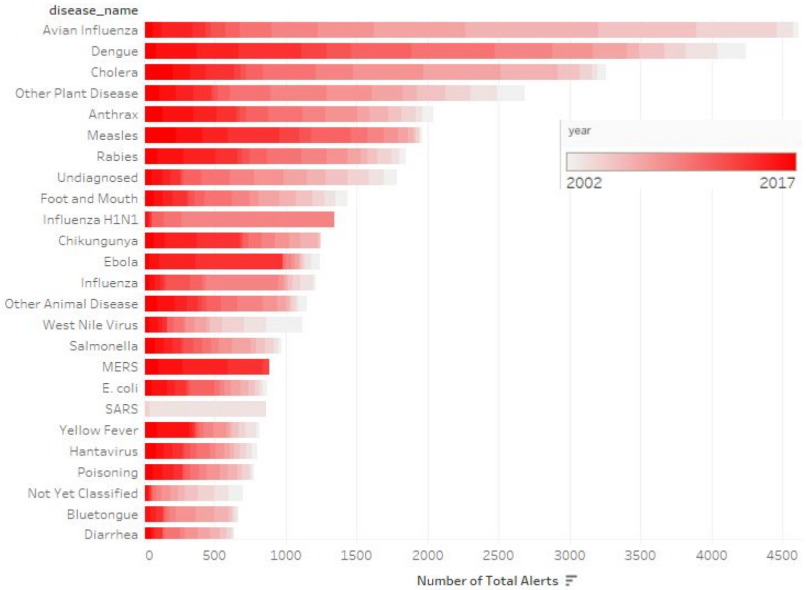
Country	Brazil
City	São Paulo
Date	10 April 2018
Species	Human
Disease	Yellow Fever
Alert content	A white spot shrimp disease outbreak in Brazil has led to a price hike of 50 per cent since the beginning of the year as local production has fallen nearly 30 per cent, officials say. According to the Brazilian newspaper ...

Dataset

The dataset includes country, time and disease for each alerts

- Total number of alerts is 36,133
- Year 2002 - 2017



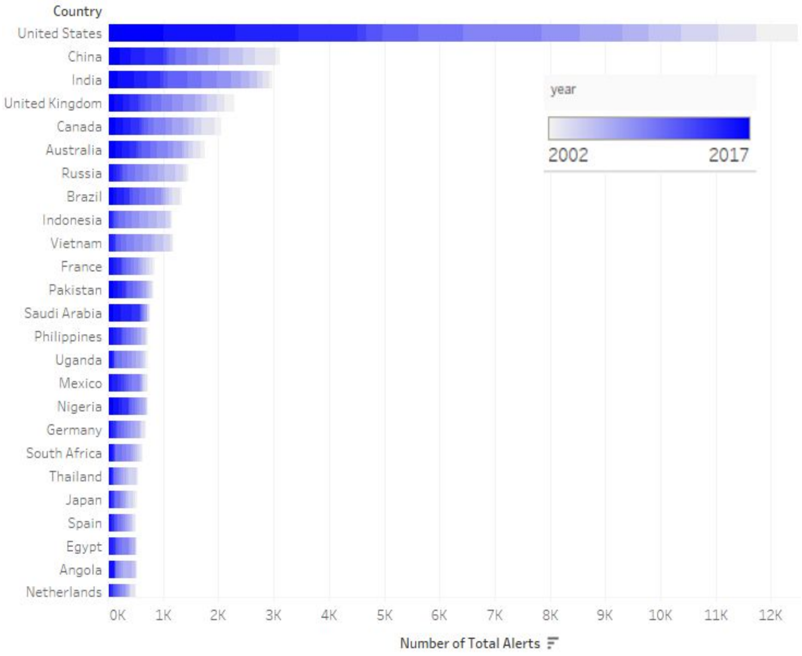


- Top 3 diseases: Avian Influenza, Dengue, Cholera
- We use Yellow Fever and African Swine Fever as samples

Dataset

The dataset we use includes country, time and disease for each alerts

- Year 2002 - 2017
- Cover 208 countries in total

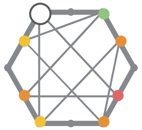


- Top 3 countries: United States, China, India

Project goal

Project goal

Giving recommendations on potential “next targets” of a particular disease given the news of an outbreak



Calculating the **probability** of an infectious disease outbreak in country B,

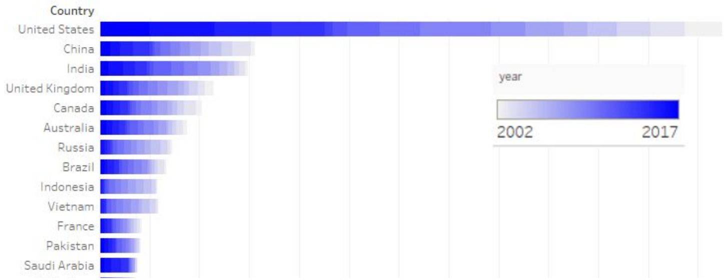
given that country A is infected by this very disease.

Value for ProMED and its network:

- **Prevention:** Setting early actions to **prevent the spread of diseases** in countries that could be infected
- **Identification: Identifying countries that are more likely to infect others**
- **Innovation:** A step further to an automatic recommendation system

Limitations

Skewed number of alerts in certain countries and unpublished alerts of common diseases can impact the model



Number of alerts by country

- 147 times more alerts for the US than the median number of alerts globally
- High number of alerts from China, India, UK and Canada

Common vs (re-) emerging diseases

- Alerts on common infectious disease are not published
- (Re-) emerging is "subjective" up to human judgement

Methodology

Methodology: Model

Modeling probabilities using vehicles of travel, time difference, country-specific factors, outbreak magnitude and environmental factors

Regressors:

1. Magnitude

Number of cases reported during an outbreak

2. Vehicles of travel

Bordering countries

Number of flights

3. Time difference

Countries with alerts further apart less likely to be related

4. Country factors

Demographic

Political stability

...

5. Environmental

Temperature

Rainfall

Dependent variable:

1 = Disease present in country B given A within period [t; t+T]

0 = Disease not present in country B given A within period [t; t+T]

Methodology: Model

Modeling probabilities using vehicles of travel, time difference, country-specific factors, outbreak magnitude and environmental factors

Regressors:

1. Magnitude

2. Vehicles of travel

3. Time difference

4. Country factors

5. Environmental

- Find and access **~ 30 external datasets** for 10+ years
- **Clean** data and fill **missing values**
- **Normalise** datasets and format as necessary

Methodology: Dropped Regressor - Magnitude

Using NLP techniques to extract the number of people infected from ProMED's alerts

1. Magnitude

Number of cases reported during an outbreak

- Used ProMED **alerts' free text information** as sources
- **Applied NLP technique to tokenize and extract noun phrases** (Date, Species, Number of Cases)
- Attempted **"Fasttext" deep learning tool for text classification**

Regressor dropped:

- Too many repetitions of cases from outbreaks in previous years or similar alerts

Methodology: Regressors

Facilitating the physical spreading of infectious diseases: vectors

2. Vehicles of travel

Connectivity impacting resilience of a country

- Number of *airports* in B accessible from A (UK/Germany);
- *Bordering* countries (US/Mexico);

Methodology: Regressors

Impact of temporal clustering

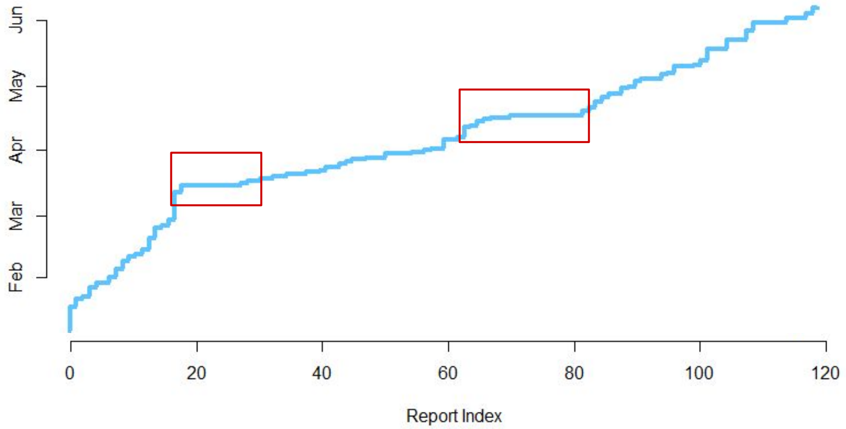
3. Time difference

Time elapsed between alerts and NOT geographical time difference between countries.

Countries with alerts submitted outside a specified time threshold less likely to be related

Capitalise on ProMED's fast response

> *clustering* of alerts



Modelling the non-infected state?

Using uniformly distributed samples to avoid clustering

Methodology: Regressors

Infectious Disease Vulnerability Index

4. Country specific factors

1. Demographic
2. Political-domestic
3. Political-international
4. Health care
5. Public Health
6. Disease Dynamics
7. Economic

- **RAND National Defense Research** Institute by expert knowledge
- Each factor composed by different **subfactors**
- Collected most of the datasets for **several years and most countries**

Methodology: Regressors

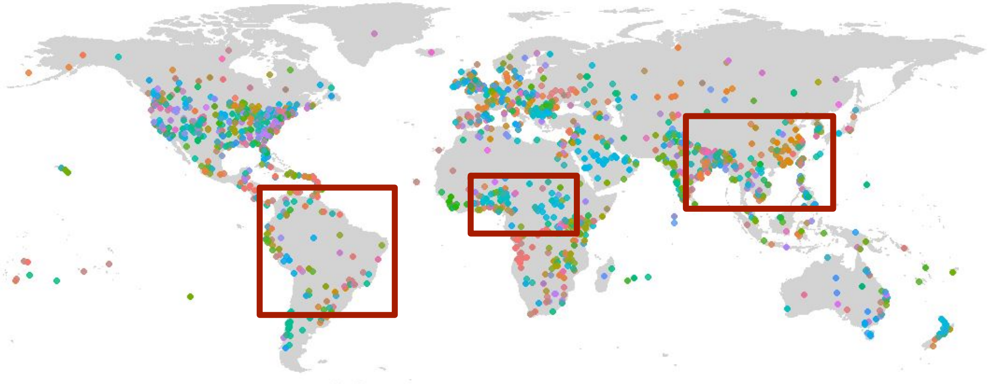
Similar environmental conditions are likely to increase the spread of an infectious disease

5. Environmental

Temperature

Rainfall

- Standard approach to reporting is very **localised** i.e. for similar environments and neglecting variations
- **Infected** country is **more likely** to spread the disease to a **similar** country**.**
- Monthly national levels of **temperature** and **rainfall (~humidity)**



Methodology: Putting It All Together - Modelling methods

Explicit vs Implicit models for supervised learning

Perfect knowledge assumption

Suppose a disease D breaks out in country A at time t. What is the **probability** that a country B is infected within a time window **[t; t+T]** given what happened in A at time t?

Class 1	Report from B within time window
Class 0	No report from B

Models used

- **Logistic regression**
  - Probit
- **Ensemble methods:**
  - Boosting (LogitBoost)
  - Random Forests

Results

Random Forest - Yellow Fever and African Swine Fever

High accuracy, but undesired results

Overall Accuracy ~ 99%

- Accuracy *“Disease present = Yes”*: ~ 99%
- Accuracy *“Disease present = No”* : ~ 99%

Issues:

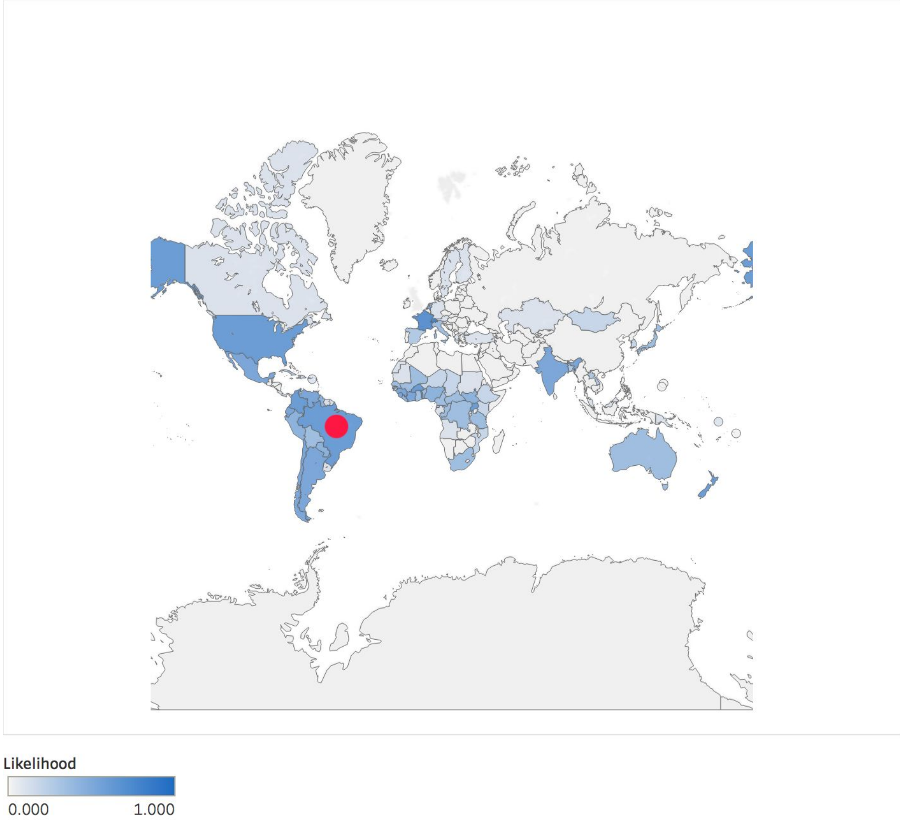
- Data **Overfitting:** cannot extrapolate
- Master dataset built on small minority of infected countries
- **Emphasis** on country B
- Modelling “commonly infected countries” rather than “spread of disease”

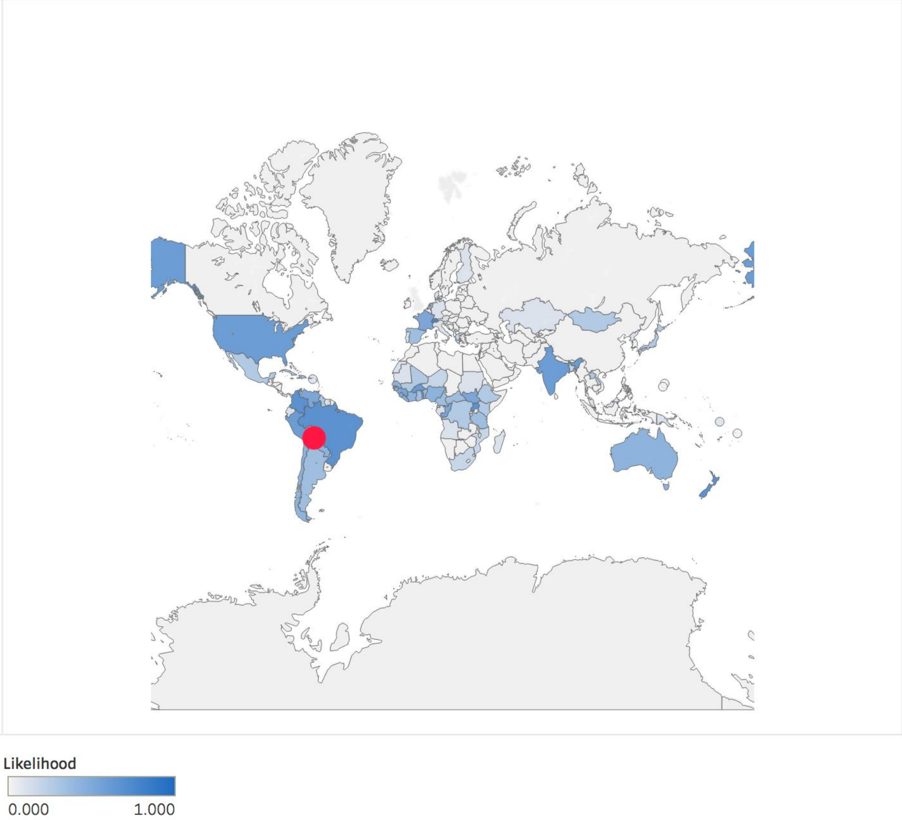
Random Forest Feature Importance (Top 10, Yellow Fever)

Feature	Feature Importance
Demographic Domain (B)	0.1671
Temperature (B)	0.1573
Rainfall (B)	0.1553
Disease dynamics (B)	0.1188
Public health (B)	0.1097
...	...

Random Forest - Yellow Fever

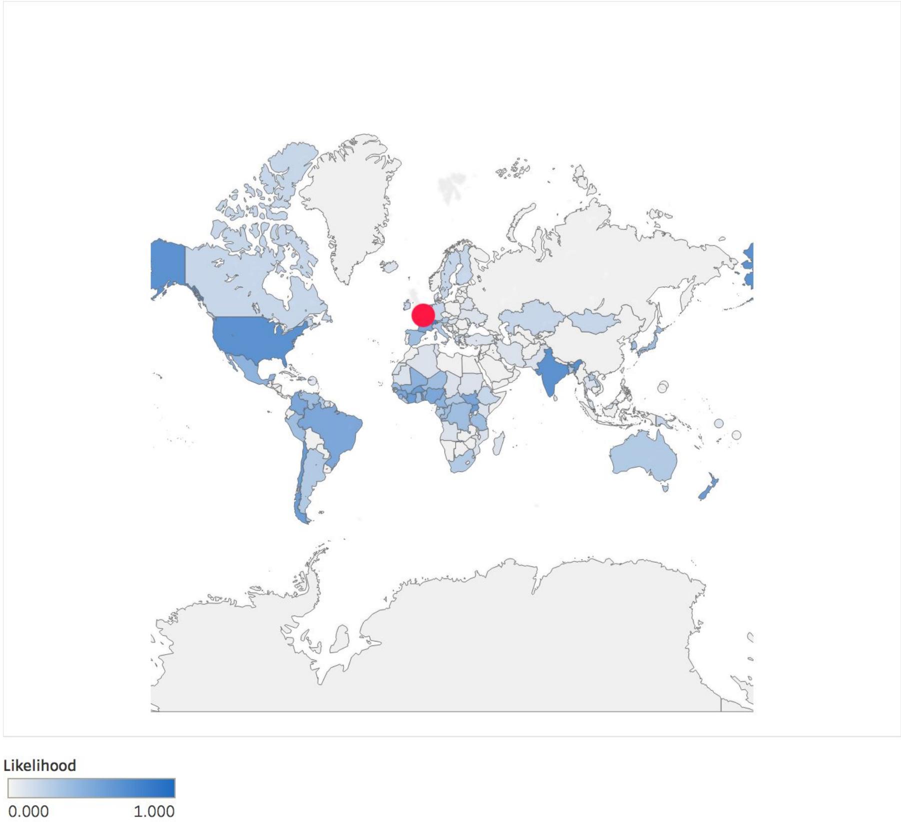
Most common infected countries always among the most likely to be infected, independently of country A

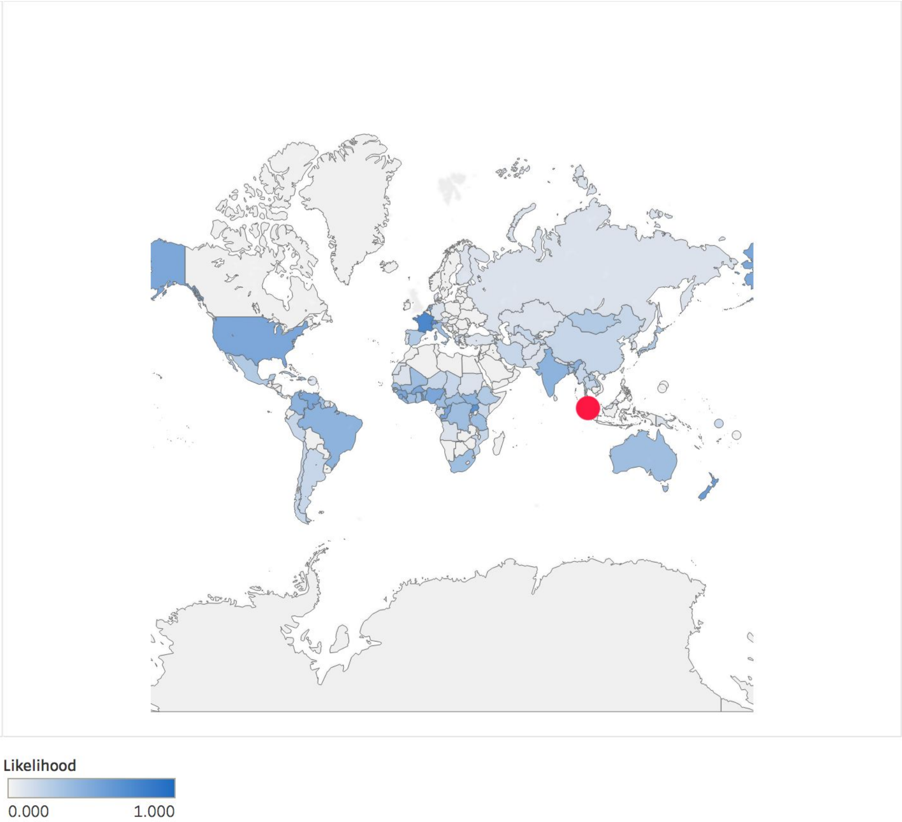




Random Forest - Yellow Fever

Most common infected countries always among the most likely to be infected, independently of country A





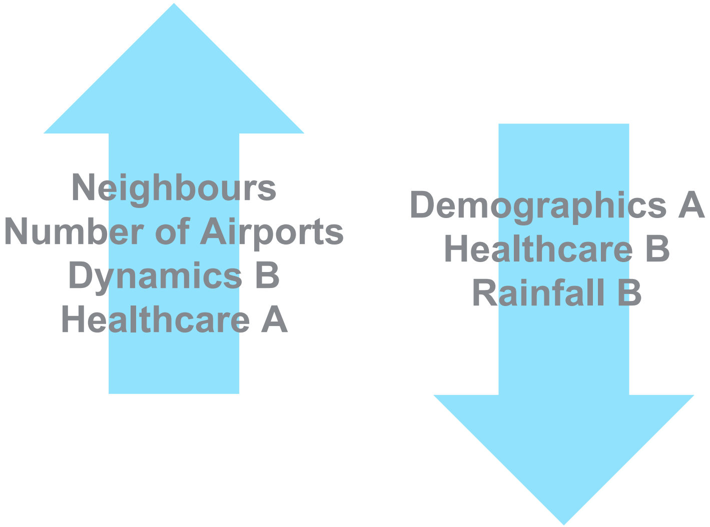
Probit Regression with Oversampling - African Swine Fever

Overall Accuracy: 72%

- Accuracy *“Disease present = Yes”*: 72%
- Accuracy *“Disease present = No”* : 68%

Marginal Effects:

- **+1 airport with A-B connecting flights** increases the likelihood of spread by **0.04**
- Being **neighbour** increases the likelihood by **0.47**
- As level of rainfall increases, likelihood decreases
- Disease dynamics (change in land use) and demographics of country B relevant (0.1 change equals to 0.08 increase in likelihood)



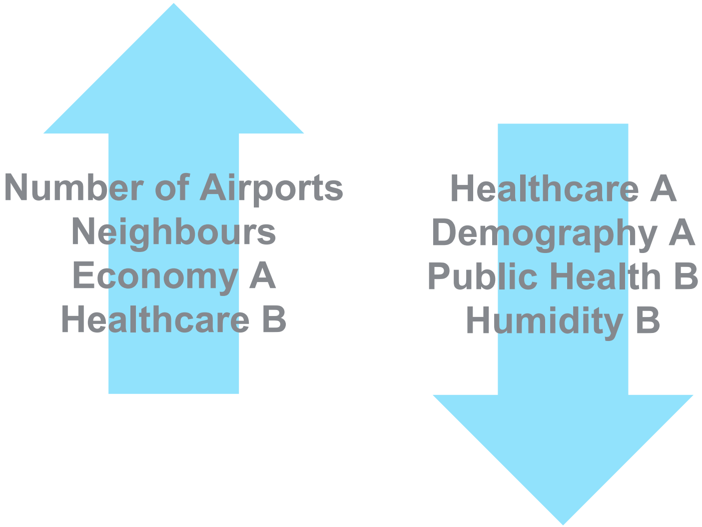
Probit Regression with Oversampling - Yellow Fever

Overall Accuracy: 72%

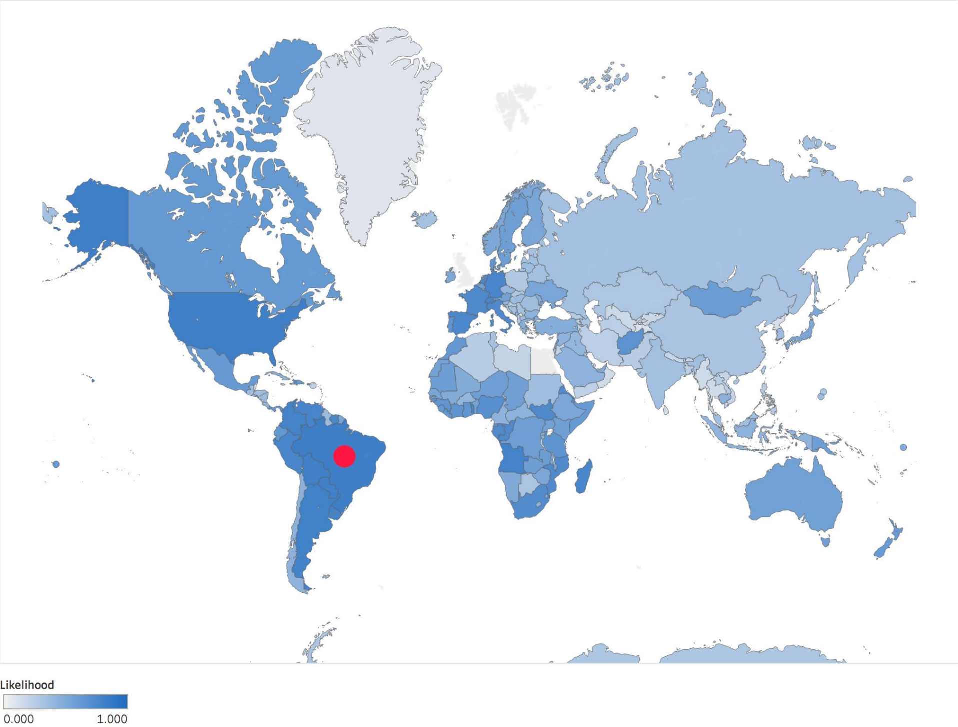
- Accuracy *“Disease present = Yes”*: 72%
- Accuracy *“Disease present = No”* : 72%

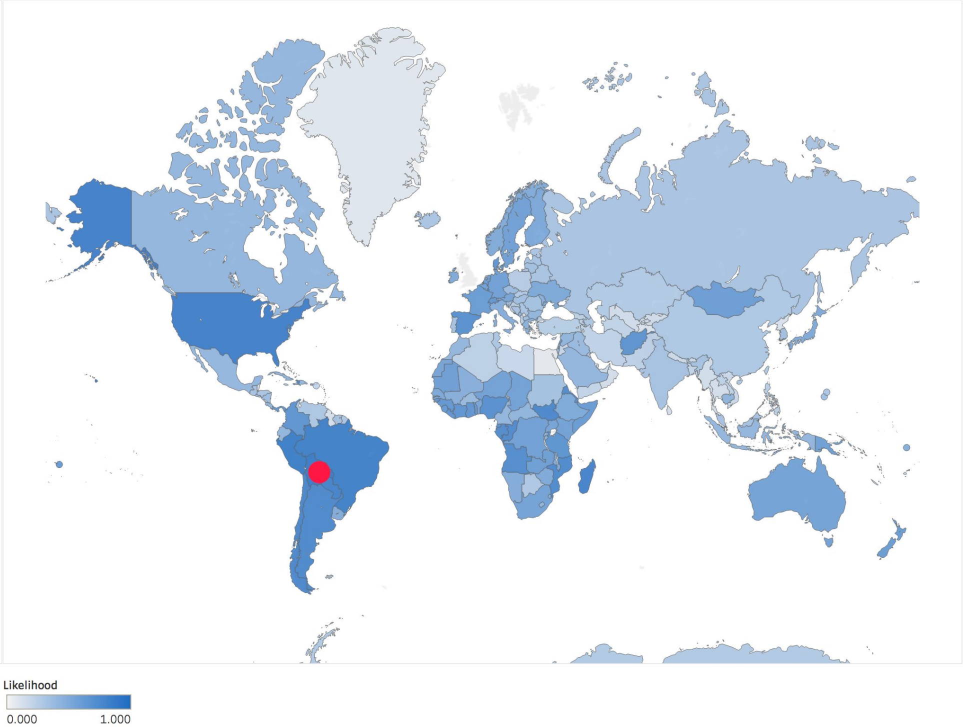
Marginal Effects:

- **+1 airport with A-B connecting flights** increases the likelihood of spread by **0.12**
- Being **neighbour** increases the likelihood by **0.31**
- Economic domain of country A very impactful (0.1 change equals to 0.28 increase in likelihood)
- Conversely, as level of rainfall increases, likelihood decreases
- Good public health decreases likelihood

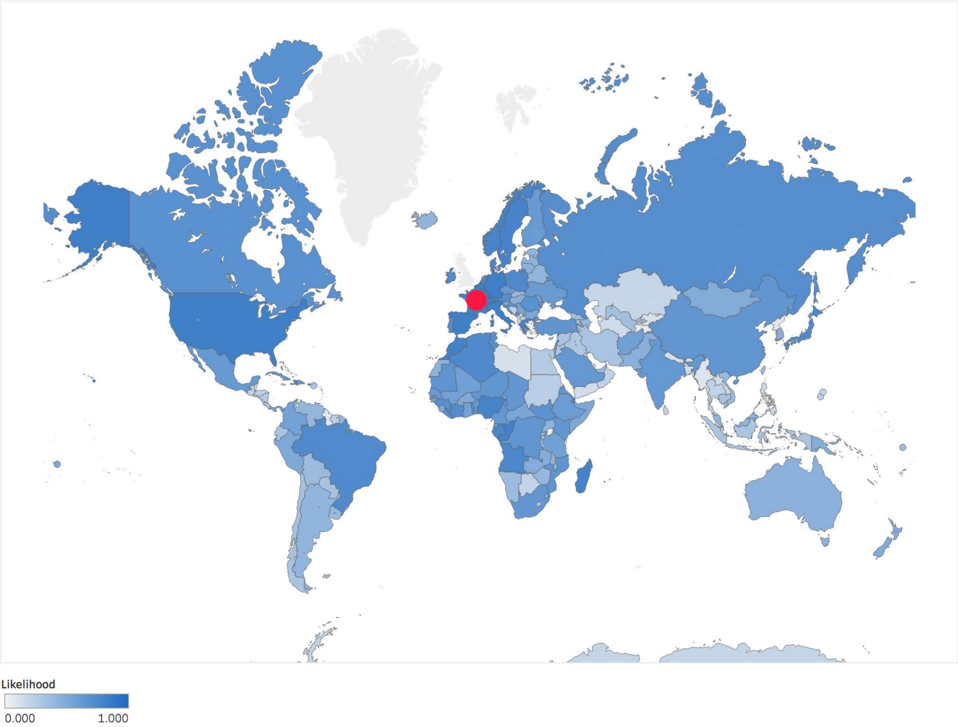


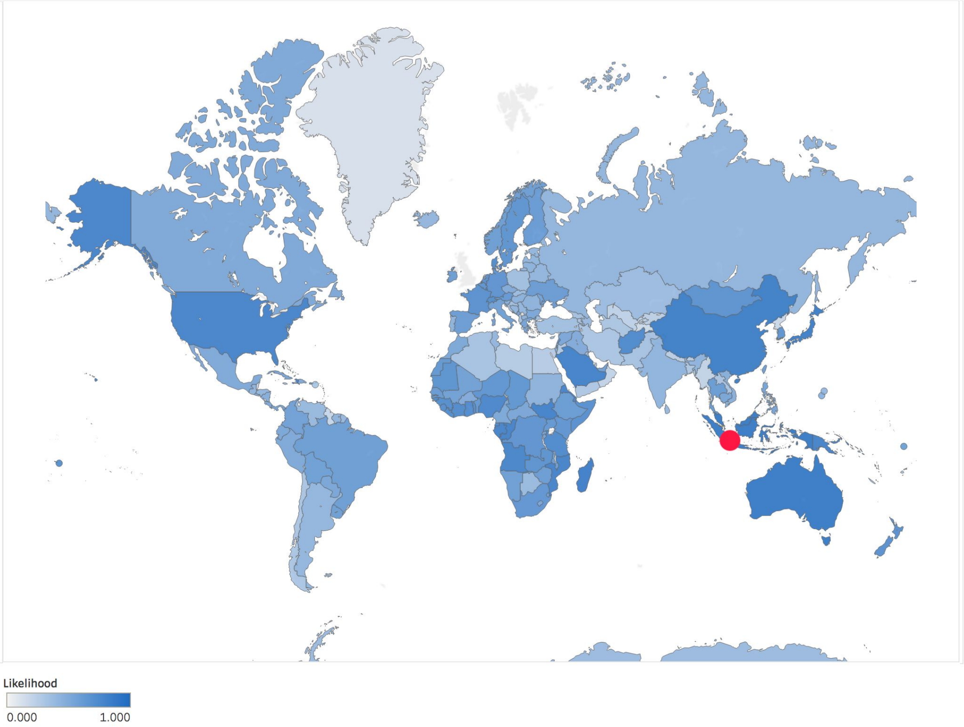
Probit Regression with Oversampling - Yellow Fever





Probit Regression with Oversampling - Yellow Fever





Infection likelihood for country B based on country A (Yellow Fever)

Out-degree, metric to identify countries that can infect many others

In-degree, metric to identify countries most vulnerable to be infected by others

Country	Out-degree
Qatar	184
Macao	179
United Arab Emirates	139
Hong Kong	114
Brunei	109
Saudi Arabia	107
Singapore	89
Trinidad and Tobago	76
Antigua and Barbuda	74
Kuwait	71

Country	In-degree
Madagascar	202
South Sudan	167
Mozambique	156
Congo	152
Haiti	152
Eritrea	144
Marshall Islands	141
Angola	140
Gabon	139
Nigeria	108

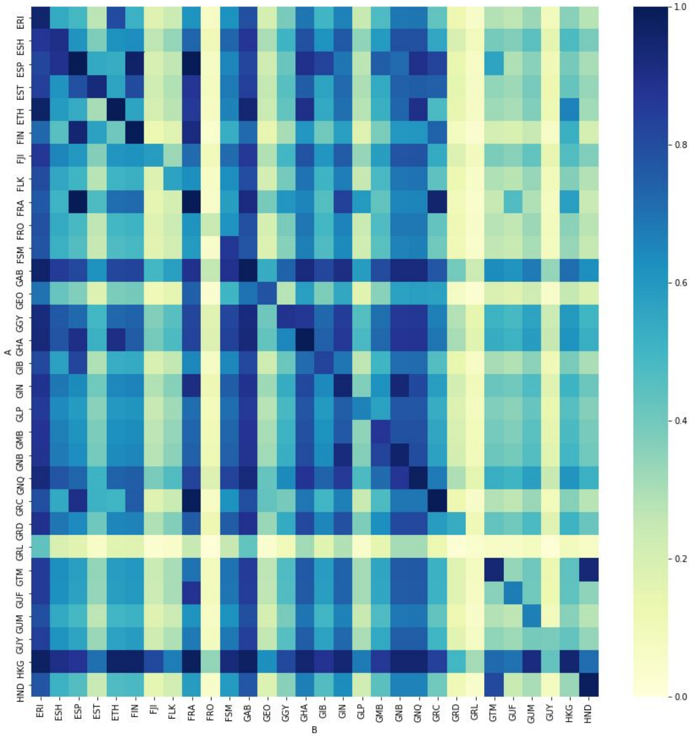
Schengen countries

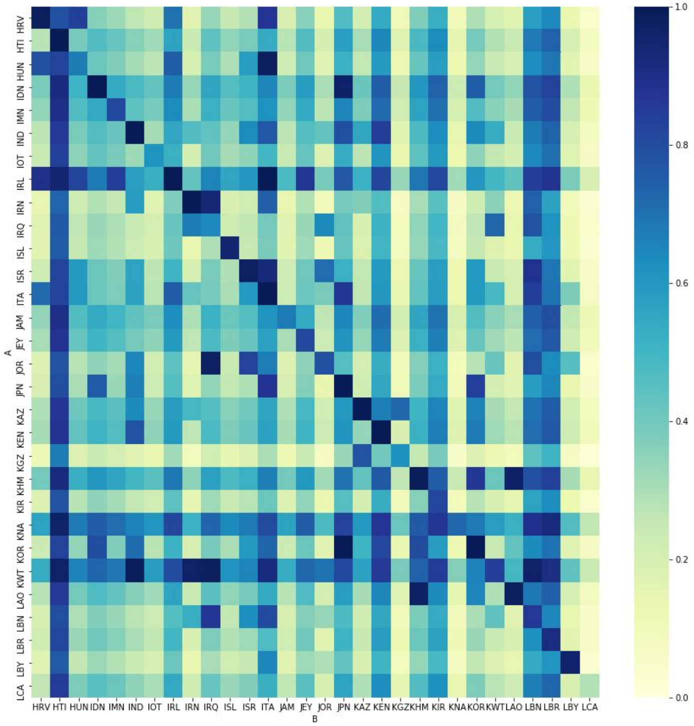
Country	In-degree
MDG	202.230044
USA	168.820683
SSD	166.671795
MOZ	156.209497
COG	151.998618
HTI	151.996562
ERI	143.535152
MHL	141.154554
AGO	140.022215
GAB	139.125168
NGA	107.669005
FRA	90.257118
DEU	85.290225
LUX	83.532799
DNK	78.257117
MAC	76.311990
CHE	74.918946
AFG	72.372336
TZA	68.618326
NLD	65.240843
GHA	61.183343
CIV	55.999092
ESP	55.777152
ITA	53.119645
GNQ	52.256237

Infection likelihood for country B based on country A (Yellow Fever)

More likely to infect many: Hong Kong

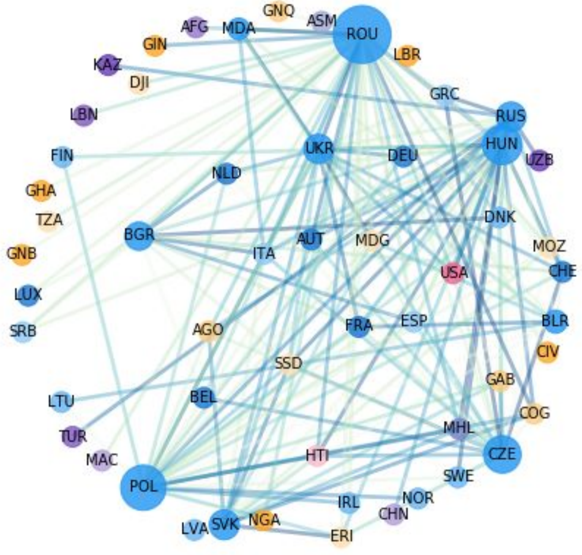
More likely to get infected: Eritrea, Gabon and Haiti



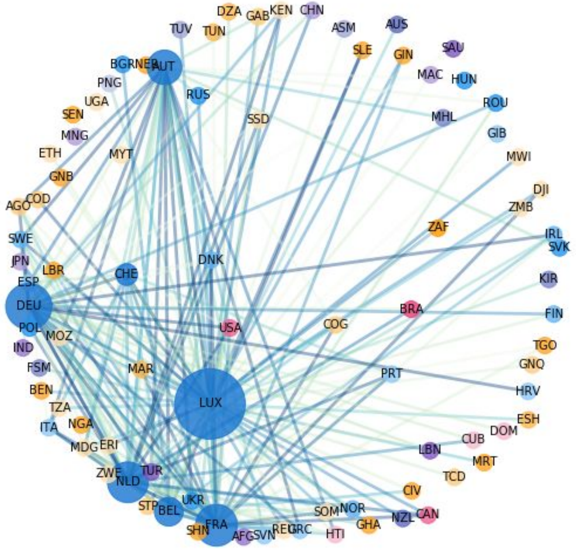


Yellow Fever likelihoods by sub-region

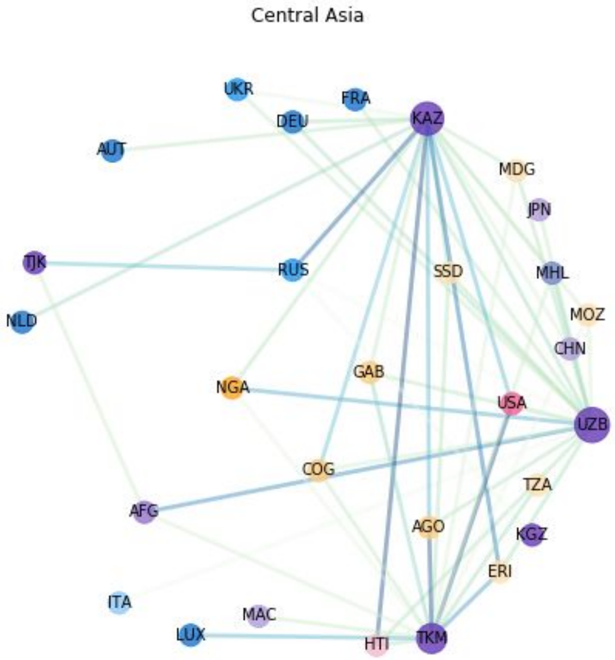
Eastern Europe

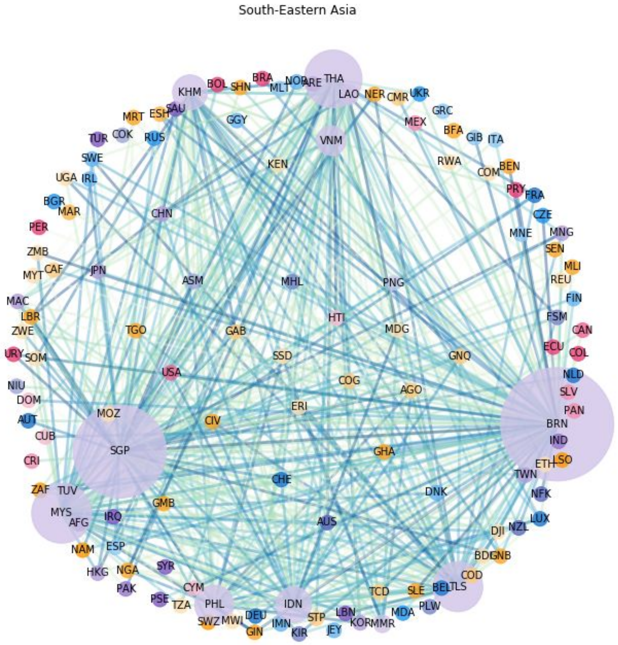


Western Europe



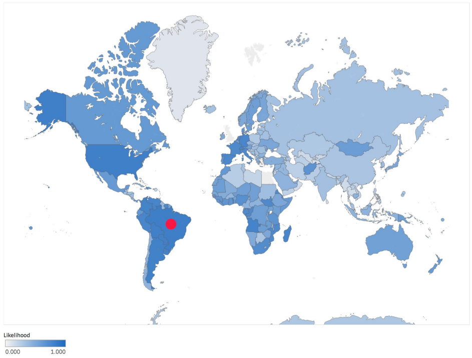
Yellow Fever likelihoods by sub-region

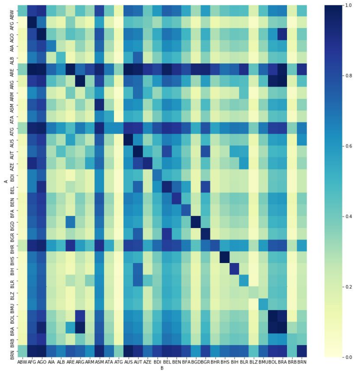


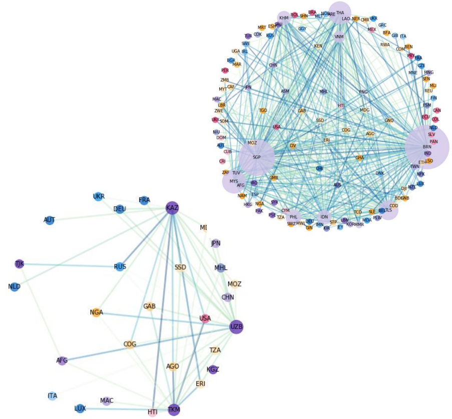


Conclusions and Recommendations

Conclusion







Modeling conditional probability of infection with Probit Regression to identify most contagious and/or vulnerable countries for infection

Some countries are more likely to infect others while some countries are more vulnerable to be infected by others

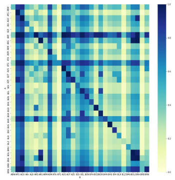
Country specific characteristics and connections affect the likelihood of infection

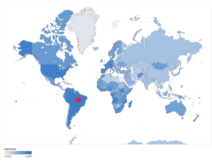
Recommendations

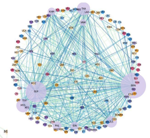
Marginal Effects:

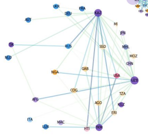
	dF/dx	Std. Err.	z	P>|z|	
NbA	0.37539433	0.00714701	52.5247	< 2.2e-16	***
factor(isNeighbour)1	0.31145697	0.00824690	37.7666	< 2.2e-16	***
TD	-0.00012021	0.00006726	-1.7872	0.07390	.
index_demo_a	-0.39088617	0.04496756	-8.6926	< 2.2e-16	***
index_demo_b	1.16619478	0.02835877	41.1229	< 2.2e-16	***
index_dynamics_a	0.21425256	0.02097053	10.2168	< 2.2e-16	***
index_dynamics_b	0.88039917	0.01529847	57.5482	< 2.2e-16	***
index_econ_a	2.78651514	0.07525970	37.0253	< 2.2e-16	***
index_econ_b	0.15471344	0.04833293	3.2010	0.00137	**
index_healthcare_a	-0.70230568	0.04256197	-16.5008	< 2.2e-16	***
index_healthcare_b	1.19429022	0.02727087	43.7936	< 2.2e-16	***
index_int_politics_a	-0.06234564	0.01453510	-4.2893	1.792e-05	***
index_int_politics_b	0.09536949	0.01162227	8.2058	2.291e-16	***
index_dom_politics_a	-0.14068785	0.01598641	-8.8005	< 2.2e-16	***
index_dom_politics_b	0.00058199	0.01122200	0.0519	0.95864	
index_pub_health_a	-0.22539755	0.02196587	-10.2613	< 2.2e-16	***
index_pub_health_b	-1.43008677	0.01468132	-97.4086	< 2.2e-16	***
index_temp_a	0.60392774	0.04531735	13.3266	< 2.2e-16	***
index_temp_b	0.30817348	0.01951738	15.7897	< 2.2e-16	***
index_rf_a	-0.06357043	0.06521921	-0.9747	0.32970	
index_rf_b	-1.22858873	0.04825550	-25.4601	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1









- Use country factors make prediction on possible changes
- Use country factors as evidence to improve certain areas of development

- Use visualisations to identify most infectious countries and more vulnerable countries
- Focus resources on such countries

Q&A

Appendix

Probit Regression

	Dependent variable:	
	IS	
	Yellow Fever	African Swine Fever
NbA	0.941***	0.425***
	(0.018)	(0.013)
factor(isNeighbour)1	0.874***	1.568***
	(0.029)	(0.040)
TD	-0.0003*	-0.006***
	(0.0002)	(0.0003)
index_demo_a	-0.980***	-2.111***
	(0.113)	(0.301)
index_demo_b	2.924***	-0.717***
	(0.071)	(0.119)
index_dynamics_a	0.537***	-0.405***
	(0.053)	(0.102)
index_dynamics_b	2.208***	2.167***
	(0.038)	(0.062)
index_econ_a	6.988***	-0.493
	(0.189)	(0.319)
index_econ_b	0.388***	1.516***
	(0.121)	(0.182)
index_healthcare_a	-1.761***	0.343*
	(0.107)	(0.176)
index_healthcare_b	2.995***	-1.306***
	(0.068)	(0.101)
index_int_politics_a	-0.156***	0.882***
	(0.036)	(0.074)
index_int_politics_b	0.239***	1.895***
	(0.029)	(0.051)
index_dom_politics_a	-0.353***	0.182*
	(0.040)	(0.099)
index_dom_politics_b	0.001	0.099**
	(0.028)	(0.046)
index_pub_health_a	-0.565***	0.363***
	(0.055)	(0.074)
index_pub_health_b	-3.586***	-0.657***
	(0.037)	(0.055)
index_temp_a	1.514***	0.164***
	(0.114)	(0.056)
index_temp_b	0.773***	-0.743***
	(0.049)	(0.068)
index_rf_a	-0.159	-0.825**
	(0.164)	(0.393)
index_rf_b	-3.081***	-5.899***
	(0.121)	(0.235)
Constant	-2.097***	0.280***
	(0.103)	(0.096)
Observations	98,432	44,308
Log Likelihood	-53,783.060	-22,468.080
Akaike Inf. Crit.	107,610.100	44,980.160
Note:		*p<0.1;* ***p<0.05;*** p<0.01

Probit Regression Results, Yellow Fever

Yellow Fever

Pred / Actual

	0	**1**
0	17,621	1,147
1	7,008	3,099

Accuracy: 0.717
Sensitivity: 0.716
Specificity: 0.723

```
Marginal Effects:
                              dF/dx   Std. Err.        z      P>|z|
NbA                      0.37539433  0.00714701  52.5247 < 2.2e-16 ***
factor(isNeighbour)1     0.31145697  0.00824690  37.7666 < 2.2e-16 ***
TD                      -0.00012021  0.00006726  -1.7872   0.07390 .
index_demo_a            -0.39088617  0.04496756  -8.6926 < 2.2e-16 ***
index_demo_b             1.16619478  0.02835877  41.1229 < 2.2e-16 ***
index_dynamics_a         0.21425256  0.02097053  10.2168 < 2.2e-16 ***
index_dynamics_b         0.88039917  0.01529847  57.5482 < 2.2e-16 ***
index_econ_a             2.78651514  0.07525970  37.0253 < 2.2e-16 ***
index_econ_b             0.15471344  0.04833293   3.2010   0.00137 **
index_healthcare_a      -0.70230568  0.04256197 -16.5008 < 2.2e-16 ***
index_healthcare_b       1.19429022  0.02727087  43.7936 < 2.2e-16 ***
index_int_politics_a    -0.06234564  0.01453510  -4.2893 1.792e-05 ***
index_int_politics_b     0.09536949  0.01162227   8.2058 2.291e-16 ***
index_dom_politics_a    -0.14068785  0.01598641  -8.8005 < 2.2e-16 ***
index_dom_politics_b     0.00058199  0.01122200   0.0519   0.95864
index_pub_health_a      -0.22539755  0.02196587 -10.2613 < 2.2e-16 ***
index_pub_health_b      -1.43008677  0.01468132 -97.4086 < 2.2e-16 ***
index_temp_a             0.60392774  0.04531735  13.3266 < 2.2e-16 ***
index_temp_b             0.30817348  0.01951738  15.7897 < 2.2e-16 ***
index_rf_a              -0.06357043  0.06521921  -0.9747   0.32970
index_rf_b              -1.22858873  0.04825550 -25.4601 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Random Forest Results, African Swine Fever

African Swine Fever

Pred / Actual

	0	**1**
0	7,935	63
1	3,145	136

Accuracy: 0.716
Sensitivity: 0.716
Specificity: 0.683

```
Marginal Effects:
                          dF/dx   Std. Err.        z     P>|z|
NbA                   0.1678169  0.0051468   32.6062 < 2.2e-16 ***
factor(isNeighbour)1  0.4758132  0.0068351   69.6131 < 2.2e-16 ***
TD                   -0.0022308  0.0001056  -21.1244 < 2.2e-16 ***
index_demo_a         -0.8338744  0.1189114   -7.0126 2.340e-12 ***
index_demo_b         -0.2831750  0.0469538   -6.0309 1.630e-09 ***
index_dynamics_a     -0.1600802  0.0402645   -3.9757 7.017e-05 ***
index_dynamics_b      0.8562037  0.0246519   34.7318 < 2.2e-16 ***
index_econ_a         -0.1946235  0.1258495   -1.5465  0.121989
index_econ_b          0.5988929  0.0720788    8.3089 < 2.2e-16 ***
index_healthcare_a    0.1353739  0.0694554    1.9491  0.051286 .
index_healthcare_b   -0.5158115  0.0400158  -12.8902 < 2.2e-16 ***
index_int_politics_a  0.3482742  0.0291784   11.9360 < 2.2e-16 ***
index_int_politics_b  0.7486514  0.0203031   36.8738 < 2.2e-16 ***
index_dom_politics_a  0.0717256  0.0391965    1.8299  0.067265 .
index_dom_politics_b  0.0391565  0.0180116    2.1740  0.029708 *
index_pub_health_a    0.1432080  0.0290749    4.9255 8.415e-07 ***
index_pub_health_b   -0.2594351  0.0218789  -11.8578 < 2.2e-16 ***
index_temp_a          0.0648726  0.0222855    2.9110  0.003603 **
index_temp_b         -0.2935242  0.0267629  -10.9676 < 2.2e-16 ***
index_rf_a           -0.3260772  0.1550819   -2.1026  0.035500 *
index_rf_b           -2.3303695  0.0932000  -25.0040 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

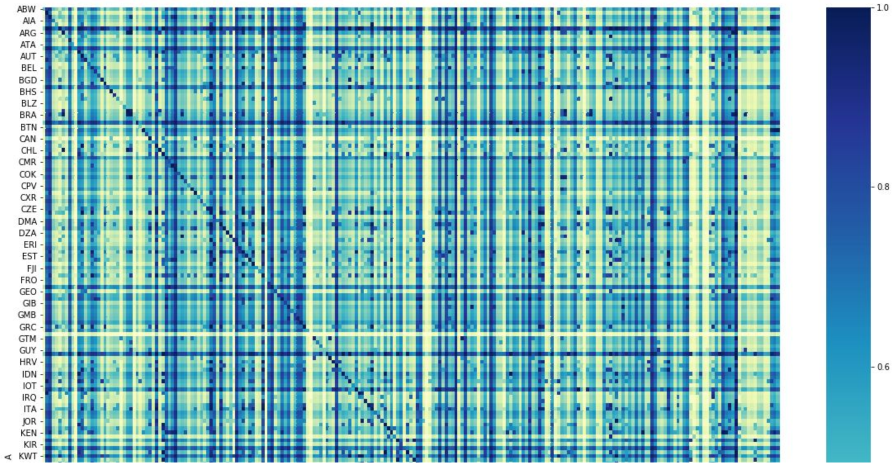
Random Forest Feature Importance, Yellow Fever

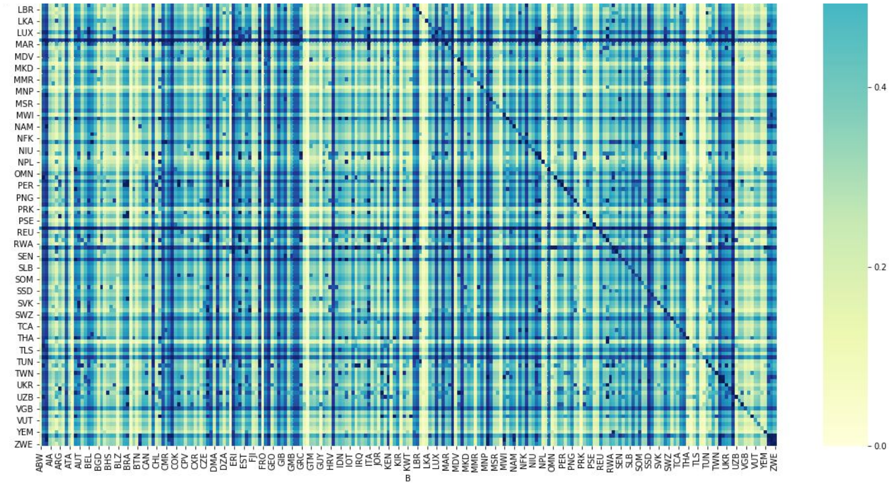
Feature	Importance
index_demo_b	0.1671
index_temp_b	0.1573
index_rf_b	0.1553
index_dynamics_b	0.1188
index_pub_health_b	0.1097
index_dom_politics_b	0.0662
NbA	0.0606
index_healthcare_b	0.0604
index_int_politics_b	0.0545
index_econ_b	0.031
index_demo_a	0.0042
index_econ_a	0.0028
index_dom_politics_a	0.0028
index_rf_a	0.0019
index_temp_a	0.0017
isNeighbour	0.0015
index_healthcare_a	0.0011
index_dynamics_a	0.0009
TD	0.0009
index_pub_health_a	0.0007
index_int_politics_a	0.0005

Random Forest Feature Importance, African Swine Fever

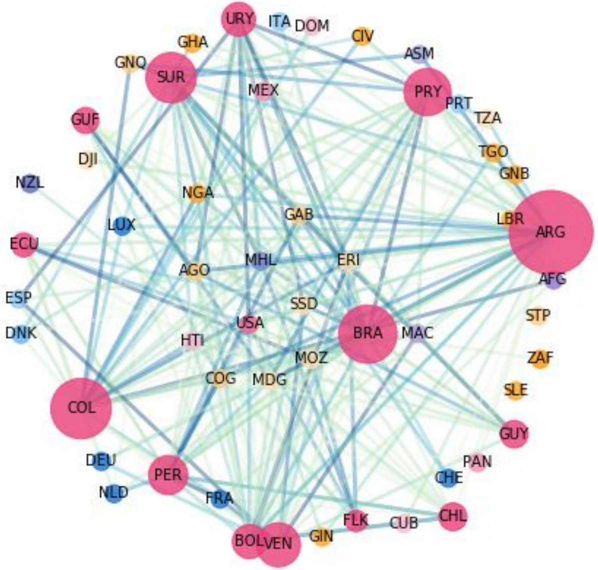
Feature	Importance
isNeighbour	0.1554
TD	0.1049
index_temp_b	0.0898
NbA	0.0767
index_rf_b	0.0684
index_pub_health_b	0.054
index_dynamics_b	0.0492
index_dom_politics_b	0.049
index_healthcare_b	0.0456
index_demo_b	0.0445
index_econ_b	0.041
index_int_politics_b	0.0392
index_temp_a	0.0279
index_pub_health_a	0.0261
index_rf_a	0.0238
index_econ_a	0.022
index_dom_politics_a	0.0198
index_healthcare_a	0.0191
index_demo_a	0.017
index_dynamics_a	0.0154
index_int_politics_a	0.0112

Infection likelihood for country B based on country A (Yellow Fever)

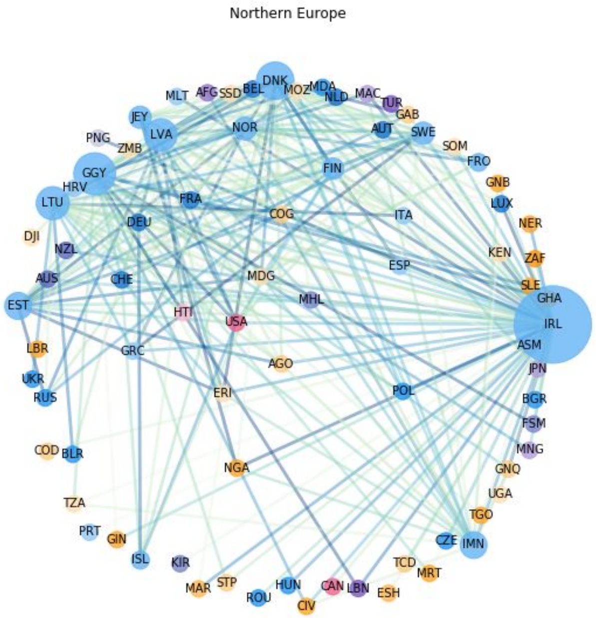


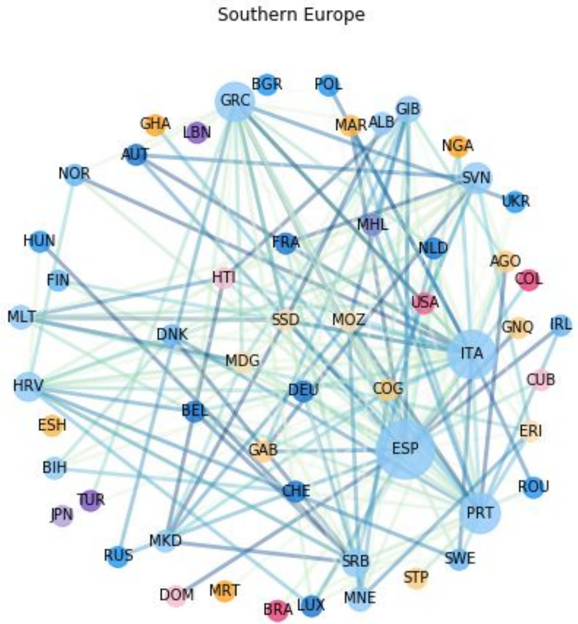


Yellow Fever likelihoods by sub-region

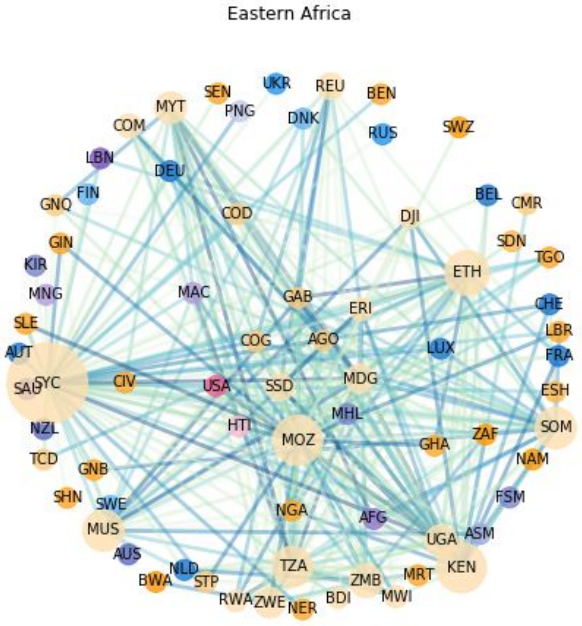


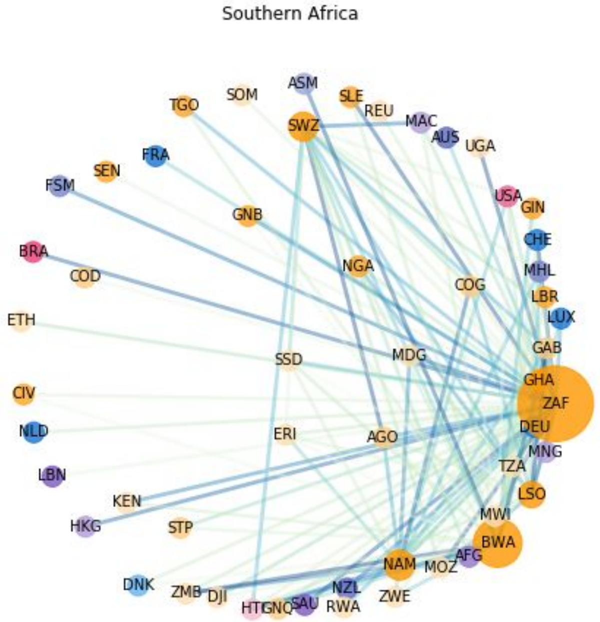
Yellow Fever likelihoods by sub-region



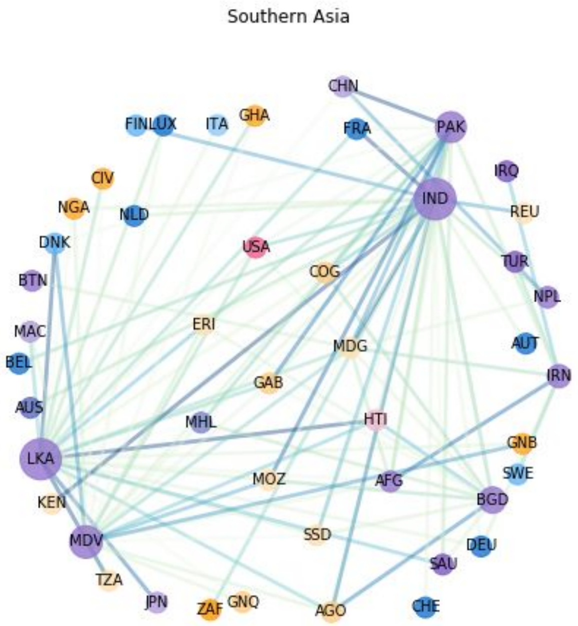


Yellow Fever likelihoods by sub-region

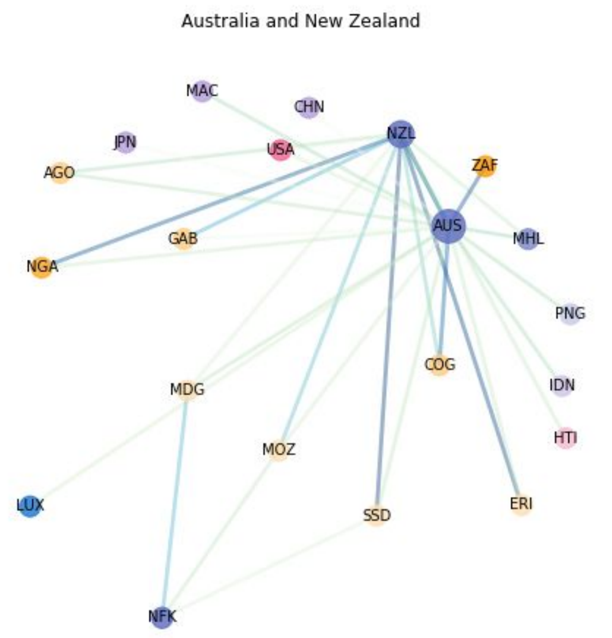




Yellow Fever likelihoods by sub-region



Yellow Fever likelihoods by sub-region

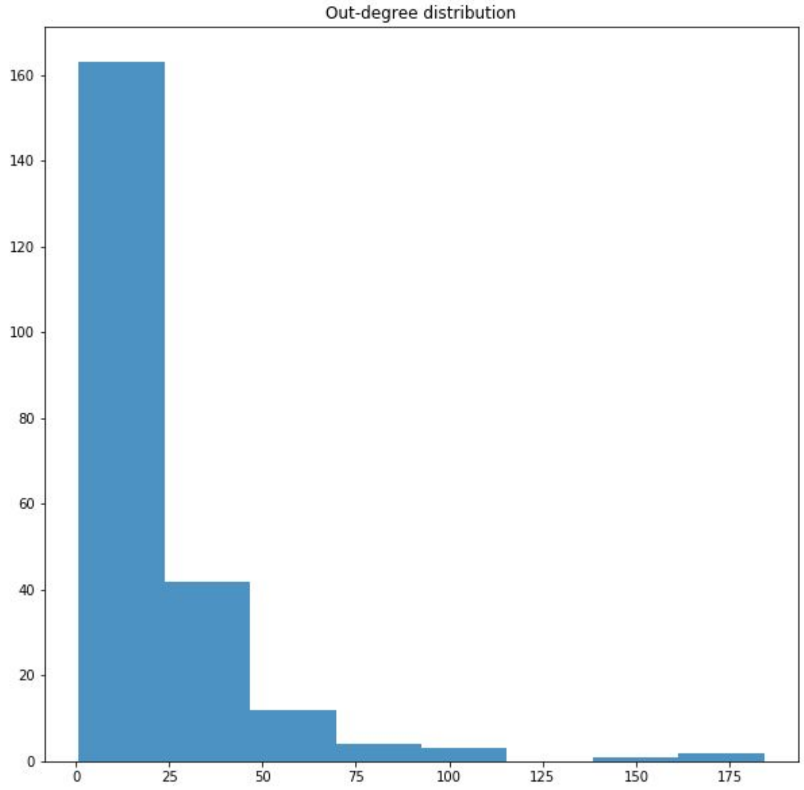


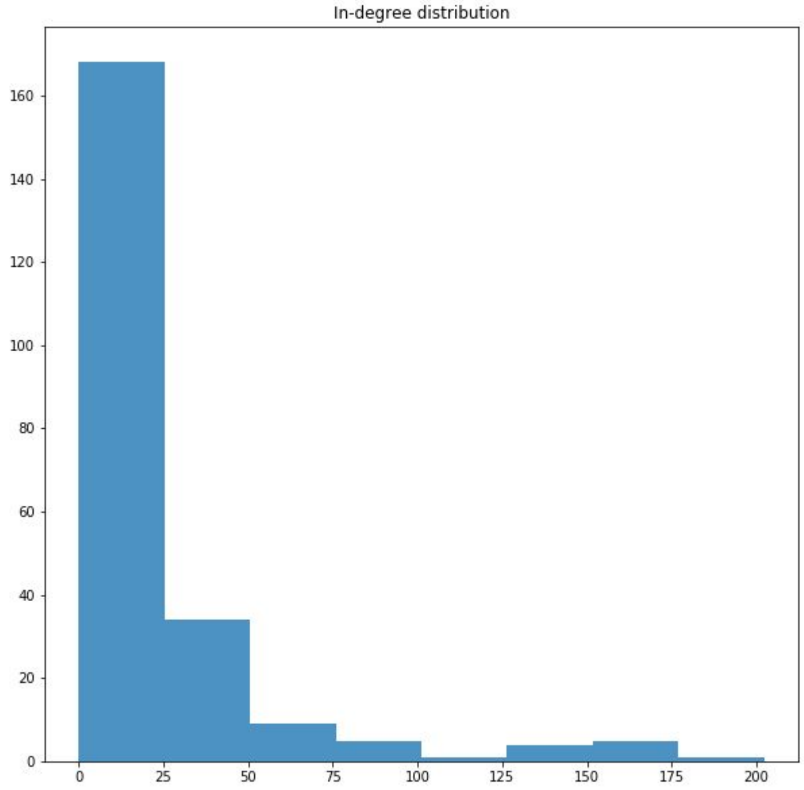
In-degree and out-degree of nodes based on likelihood >= 0.8

Country	Out-degree
QAT	184.246879
MAC	178.880751
ARE	139.128844
HKG	114.988829
BRN	109.079793
SAU	107.162042
SGP	89.349211
TTO	76.833306
ATG	74.990204
KWT	71.124087
KNA	66.237993
IRL	64.655876
GAB	63.493150
VGB	62.506483
PAN	62.410531
LUX	58.436661
BHR	56.972756
MYS	56.516199
THA	54.312024
SYC	53.278830

Country	In-degree
MDG	202.230044
USA	168.820683
SSD	166.671795
MOZ	156.209497
COG	151.998618
HTI	151.996562
ERI	143.535152
MHL	141.154554
AGO	140.022215
GAB	139.125168
NGA	107.669005
FRA	90.257118
DEU	85.290225
LUX	83.532799
DNK	78.257117
MAC	76.311990
CHE	74.918946
AFG	72.372336
TZA	68.618326
NLD	65.240843
GHA	61.183343
CIV	55.999092
ESP	55.777152
ITA	53.119645
GNQ	52.256237

In-degree and out-degree of nodes based on likelihood >= 0.8





Probit regression, Yellow Fever

Marginal effects

```
Marginal Effects:
                          dF/dx   Std. Err.        z     P>|z|
NbA                  0.37539433  0.00714701  52.5247 < 2.2e-16 ***
factor(isNeighbour)1 0.31145697  0.00824690  37.7666 < 2.2e-16 ***
TD                  -0.00012021  0.00006726  -1.7872   0.07390 .
index_demo_a        -0.39088617  0.04496756  -8.6926 < 2.2e-16 ***
index_demo_b         1.16619478  0.02835877  41.1229 < 2.2e-16 ***
index_dynamics_a     0.21425256  0.02097053  10.2168 < 2.2e-16 ***
index_dynamics_b     0.88039917  0.01529847  57.5482 < 2.2e-16 ***
index_econ_a         2.78651514  0.07525970  37.0253 < 2.2e-16 ***
index_econ_b         0.15471344  0.04833293   3.2010   0.00137 **
index_healthcare_a  -0.70230568  0.04256197 -16.5008 < 2.2e-16 ***
index_healthcare_b   1.19429022  0.02727087  43.7936 < 2.2e-16 ***
index_int_politics_a -0.06234564 0.01453510  -4.2893 1.792e-05 ***
index_int_politics_b  0.09536949 0.01162227   8.2058 2.291e-16 ***
index_dom_politics_a -0.14068785 0.01598641  -8.8005 < 2.2e-16 ***
index_dom_politics_b  0.00058199 0.01122200   0.0519   0.95864
index_pub_health_a  -0.22539755  0.02196587 -10.2613 < 2.2e-16 ***
index_pub_health_b  -1.43008677  0.01468132 -97.4086 < 2.2e-16 ***
index_temp_a         0.60392774  0.04531735  13.3266 < 2.2e-16 ***
index_temp_b         0.30817348  0.01951738  15.7897 < 2.2e-16 ***
index_rf_a          -0.06357043  0.06521921  -0.9747   0.32970
index_rf_b          -1.22858873  0.04825550 -25.4601 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Probit regression, African Swine Fever

Marginal effects

```
Marginal Effects:
                           dF/dx  Std. Err.         z     P>|z|
NbA                    0.1678169  0.0051468   32.6062 < 2.2e-16 ***
factor(isNeighbour)1   0.4758132  0.0068351   69.6131 < 2.2e-16 ***
TD                    -0.0022308  0.0001056  -21.1244 < 2.2e-16 ***
index_demo_a          -0.8338744  0.1189114   -7.0126 2.340e-12 ***
index_demo_b          -0.2831750  0.0469538   -6.0309 1.630e-09 ***
index_dynamics_a      -0.1600802  0.0402645   -3.9757 7.017e-05 ***
index_dynamics_b       0.8562037  0.0246519   34.7318 < 2.2e-16 ***
index_econ_a          -0.1946235  0.1258495   -1.5465  0.121989
index_econ_b           0.5988929  0.0720788    8.3089 < 2.2e-16 ***
index_healthcare_a     0.1353739  0.0694554    1.9491  0.051286 .
index_healthcare_b    -0.5158115  0.0400158  -12.8902 < 2.2e-16 ***
index_int_politics_a   0.3482742  0.0291784   11.9360 < 2.2e-16 ***
index_int_politics_b   0.7486514  0.0203031   36.8738 < 2.2e-16 ***
index_dom_politics_a   0.0717256  0.0391965    1.8299  0.067265 .
index_dom_politics_b   0.0391565  0.0180116    2.1740  0.029708 *
index_pub_health_a     0.1432080  0.0290749    4.9255 8.415e-07 ***
index_pub_health_b    -0.2594351  0.0218789  -11.8578 < 2.2e-16 ***
index_temp_a           0.0648726  0.0222855    2.9110  0.003603 **
index_temp_b          -0.2935242  0.0267629  -10.9676 < 2.2e-16 ***
index_rf_a            -0.3260772  0.1550819   -2.1026  0.035500 *
index_rf_b            -2.3303695  0.0932000  -25.0040 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```